Overlooked Covariates in Metabolite Abundances: Systematically Quantifying the Information Overlap between Gene Expression and Metabolism across multiple cancer types

Introduction

Reprogrammed Metabolism

- Essential for tumor proliferation and metastasis
- Assists in immune evasion

Unable to effectively computationally predict or experimentally measure metabolism

There has been no systematic quantification of the predictive power of gene expression, nor of the potential of other metadata as metabolic regulators

Methodology

Sourcing relevant and applicable datasets

- Cancer Cell Line Encyclopedia (CCLE)
- Cancer Atlas of Metabolic Profiles (CAMP)

Exploring the effect of covariates on metabolomic variance

Developing standardized preprocessing and curation

Identifying information overlap of gene expression and covariates

Distance Correlation (DC)

Partial Distance Correlation (pDC)

Novel non-linear metric of information overlap

Expresses proportion of DC related to covariates

Results

<table>
<thead>
<tr>
<th>PC</th>
<th>Associated Covariate/s</th>
<th>% of Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tissue of Origin</td>
<td>20.2%</td>
</tr>
<tr>
<td>2</td>
<td>Growth Rate, Mutation Rate</td>
<td>8.2%</td>
</tr>
<tr>
<td>3</td>
<td>Growth Rate, Mutation Rate</td>
<td>7.1%</td>
</tr>
<tr>
<td>4</td>
<td>Extracellular Environment</td>
<td>5.0%</td>
</tr>
<tr>
<td>5</td>
<td>Extracellular Environment</td>
<td>4.7%</td>
</tr>
</tbody>
</table>

Cancer Cell Line Encyclopedia

Sampling testing reveals more metabolic information is being associated with random genes than metabolic genes

Tissue of origin masks the true information overlap between gene expression and metabolism

Only ~40% of metabolic variance is associated with gene expression

External covariates have significant effect and predictive power with cancer metabolism

References


[Notes: (*) Created with Biorender.co]